ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ
INFORMATION TECHNOLOGY, COMPUTER SCIENCE, AND MANAGEMENT



УДК 519.211 DOI 10.12737/16076

О применимости формулы Байеса*

А. И. Долгов[1**]
[1]Акционерное общество «Конструкторское бюро по радиоконтролю систем управления, навигации и связи», г. Ростов-на-Дону, Российская Федерация

On applicability of Bayes' formula***

A. I. Dolgov[1**]
[1]«Design bureau on monitoring of control, navigation and communication systems» JSC, Rostov-on-Don, Russian Federation

Предметом данного исследования является формула Байеса. Цель настоящей работы — анализ и расширение области применения формулы. Первоочередной задачей представляется изучение публикаций, посвященных указанной проблеме, позволившее выявить недостатки применения формулы Байеса, приводящие к некорректным результатам. Следующая задача — построение модификаций формулы Байеса, обеспечивающих учет различных одиночных свидетельств с получением корректных результатов. И, наконец, на примере конкретных исходных данных сравниваются некорректные результаты, получаемые с применением формулы Байеса, и корректные результаты, вычисляемые с помощью предлагаемых модификаций. При проведении исследования использованы два метода. Во-первых, проведен анализ принципов построения известных выражений, применяемых для записи формулы Байеса и ее модификаций. Во-вторых, выполнена сравнительная оценка результатов (в том числе количественная). Предлагаемые модификации обеспечивают более широкое применение формулы Байеса в теории и на практике, в том числе при решении прикладных задач.

Ключевые слова: условные вероятности, несовместные гипотезы, совместимые и несовместимые свидетельства, нормирование.

Bayes' formula is the research subject. The work objective is to analyze the formula application and widen the scope of its applicability. The first-priority problem includes the identification of the Bayes' formula disadvantages based on the study of the relevant publications leading to incorrect results. The next task is to construct the Bayes' formula modifications to provide an accounting of various single indications to obtain correct results. And finally, the incorrect results obtained with the application of Bayes' formula are compared to the correct results calculated with the use of the proposed formula modifications by the example of the specific initial data. Two methods are used in studies. First, the analysis of the principles of constructing the known expressions used to record the Bayesian formula and its modifications is conducted. Secondly, a comparative evaluation of the results (including the quantitative one) is performed. The proposed modifications provide a wider application of Bayes' formula both in theory and practice including the solution of the applied problems.

Keywords: conditional probabilities, inconsistent hypotheses, compatible and incompatible indications, normalizing.

Введение. Формула Байеса находит все более широкое применение в теории и практике [1–10], в том числе при решении прикладных задач с помощью вычислительной техники [5 и 6]. Использование взаимно независимых вычислительных процедур позволяет особенно эффективно применять данную формулу при решении задач на многопроцессорных вычислительных системах [9], так как в этом случае параллельная реализация выполняется на уровне общей схемы, и при добавлении очередного алгоритма или класса задач нет необходимости повторно проводить работу по распараллеливанию.

Предметом данного исследования является применимость формулы Байеса для сравнительной оценки апостериорных условных вероятностей несовместных гипотез при различных одиночных свидетельствах. Как показывает анализ, в таких случаях сравниваются нормированные вероятности несовместных комбинированных событий, принадле-

*Работа выполнена в рамках инициативной НИР.
**E-mail: dolgov-ai@yandex.ru
***The research is done within the frame of the independent R&D.

жащих разным полным группам событий [7–9]. При этом сравниваемые результаты оказываются неадекватными реальным статистическим данным. Это обусловлено следующими факторами:
— используется некорректное нормирование [10];
— не принимается во внимание наличие или отсутствие пересечений учитываемых свидетельств.

С целью устранения обнаруженных недостатков выявляются случаи применимости формулы Байеса. Если же указанная формула неприменима, решается задача построения ее модификации, обеспечивающей учет различных одиночных свидетельств с получением корректных результатов. На примере конкретных исходных данных выполнена сравнительная оценка результатов:
— некорректных — получаемых с использованием формулы Байеса;
— корректных — вычисляемых с помощью предлагаемой модификации.

Исходные положения. В основу излагаемых далее утверждений положим принцип сохранения отношений вероятностей: «Корректная обработка вероятностей событий осуществима лишь при нормировании с применением одного общего нормирующего делителя, обеспечивающего равенство отношений нормированных вероятностей отношениям соответствующих им нормируемых вероятностей» [10]. Данный принцип представляет субъективную основу теории вероятностей, однако не отражается должным образом в современной учебной и научно-технической литературе.

При нарушении указанного принципа искажаются сведения о степени возможности рассматриваемых событий. Получаемые на основе искаженных сведений результаты и принимаемые решения оказываются неадекватными реальным статистическим данным.

В предлагаемой статье будут использованы следующие понятия:
— элементарное событие — событие, не делимое на элементы;
— комбинированное событие — событие, представляющее то или иное сочетание элементарных событий;
— совместимые события — события, которые в одних случаях сравнительной оценки их вероятностей могут быть несовместными, а других случаях совместными;
— несовместимые события — события, которые во всех случаях являются несовместными.

В качестве элементарных событий далее будут рассматриваться подтверждаемые гипотезы и свидетельства, а в качестве комбинированных — только произведения элементарных событий.

Согласно теореме умножения вероятностей, вероятность $P(H_k E)$ произведения элементарных событий H_k и E вычисляется в виде произведения вероятностей $P(H_k E) = P(E) P(H_k|E)$ [2]. В связи с этим формула Байеса часто записывается в виде $P(H_k|E) = \dfrac{P(H_k E)}{P(E)}$, описывающем определение апостериорных условных вероятностей $P(H_k|E)$ гипотез H_k $(k = 1,\ldots,n)$ на основе нормирования априорных вероятностей $P(H_k E)$ учитываемых комбинированных несовместимых событий $H_k E$. Каждое из таких событий представляет произведение, сомножителями которого являются одна из рассматриваемых гипотез и одно учитываемое свидетельство. При этом все рассматриваемые события $H_k E$ $(k = 1,\ldots,n)$ образуют полную группу $\bigcup_{k=1}^{n} H_k E$ несовместимых комбинированных событий, в связи с чем их вероятности $P(H_k E)$ должны быть нормированы с учетом формулы полной вероятности [2], согласно которой $P(E) = \sum_{k=1}^{n} P(H_k) P(E|H_k)$. Поэтому формула Байеса чаще всего записывается в наиболее употребляемом виде:

$$P(H_k|E) = \frac{P(H_k)\, P(E|H_k)}{\sum_{k=1}^{n} P(H_k)\, P(E|H_k)}. \tag{1}$$

Далее измененная формула, в которой учитывается более чем одно свидетельство, будет называться модификацией формулы Байеса.

Анализ особенностей построения формулы Байеса, нацеленного на решение прикладных задач, а также примеры ее практического применения позволяют сделать важный вывод относительно выбора полной группы сравниваемых по степени возможности комбинированных событий (каждое из которых является произведением двух элементарных событий — одной из гипотез и учитываемого свидетельства). Такой выбор осуществляется субъективно лицом, принимающим решение, на основе объективных исходных данных, присущих типовым условиям обстановки: виды и количество оцениваемых гипотез и конкретно учитываемое свидетельство.

Несравниваемые вероятности гипотез при одиночных несовместимых свидетельствах. Формула Байеса традиционно применяется в случае определения не сравниваемых по степени возможности апостериорных условных веро-

ятностей гипотез H_k при одиночных несовместимых свидетельствах E_i, каждое из которых может «появиться только в комбинации с какой-либо из этих гипотез» [2]. При этом выбираются полные группы $\bigcup_{k=1}^{n} H_k E_i$ комбинированных событий в виде произведений, сомножителями которых являются одно из свидетельств E_i (i=1,…,m) и одна из n рассматриваемых гипотез.

Формула Байеса применяется для сравнительной оценки вероятностей комбинированных событий каждой такой полной группы, отличающейся от других полных групп не только учитываемым свидетельством E_i, но и в общем случае видами гипотез H_k и (или) их количеством n (см., например, [2])

$$P(H_k|E_i) = \frac{P(H_k)\ P(E_i|H_k)}{\sum\limits_{k=1}^{n} P(H_k)\ P(E_i|H_k)}.$$

В частном случае при n = 2

$$P(H_k|E_i) = \frac{P(H_k) P(E_i|H_k)}{\sum\limits_{k=1}^{2} P(H_k) P(E_i|H_k)}.$$

и получаемые результаты являются правильными, ввиду соблюдения принципа сохранения отношений вероятностей:

$$\frac{P(H_1|E_i)}{P(H_2|E_i)} = \frac{P(H_1) P(E_i|H_1)}{\sum\limits_{k=1}^{2} P(H_k) P(E_i|H_k)} \Bigg/ \frac{P(H_2)\ P(E_i|H_2)}{\sum\limits_{k=1}^{2} P(H_k)\ P(E_i|H_k)} = \frac{P(H_1)\ P(E_i|H_1)}{P(H_2)\ P(E_i|H_2)}.$$

Субъективность выбора полной группы сравниваемых по степени возможности комбинированных событий (с теми или иными изменяемыми элементарными событиями) позволяет выбрать полную группу событий $\bigcup_{k=1}^{n} H_k\ \overline{E}_i$ с отрицанием элементарного события $\overline{E}_i$ ([6–8]) и записать формулу Байеса (i = 1,…,m) так:

$$P(H_k|\overline{E}_i) = \frac{P(H_k) P(\overline{E}_i|H_k)}{\sum\limits_{k=1}^{n} P(H_k) P(\overline{E}_i|H_k)}.$$

Такая формула также применима и дает возможность получить правильные результаты, если вычисляемые нормированные вероятности сравниваются при различных рассматриваемых гипотезах, но не при различных свидетельствах.

Сравниваемые вероятности гипотез при одиночных несовместимых свидетельствах. Судя по известным публикациям (за исключением работ автора излагаемого материала — например, [9] и [10]), формула Байеса широко применяется для сравнительной оценки апостериорных условных вероятностей гипотез при различных одиночных свидетельствах. При этом не уделяется внимание следующему факту. В указанных случаях сравниваются нормируемые вероятности несовместных (несовместимых) комбинированных событий, принадлежащих разным полным группам событий. Однако в данном случае формула Байеса неприменима, так как сравниваются не входящие в одну полную группу комбинированные события, нормирование вероятностей которых осуществляется с использованием разных нормирующих делителей. Нормированные вероятности несовместных (несовместимых) комбинированных событий можно сравнивать только в том случае, если они принадлежат одной и той же полной группе событий и нормированы с использованием общего делителя, равного сумме вероятностей всех нормируемых событий, входящих в полную группу.

В общем случае в качестве несовместимых свидетельств могут рассматриваться:

— два свидетельства (например, свидетельство и его отрицание);

— три свидетельства (к примеру, в игровой ситуации выигрыш, проигрыш и ничья);

— четыре свидетельства (в частности, в спорте выигрыш, проигрыш, ничья и переигровка) и т. д.

Рассмотрим довольно простой пример (соответствующий примеру, приведенному в [8]) применения формулы Байеса для определения апостериорных условных вероятностей гипотезы H_k при двух несовместимых событиях в виде свидетельства A_j и его отрицания $\overline{A}_j$:

$$P(H_k | A_j) = \frac{P(H_k A_j)}{P(A_j)} = \frac{P(H_k)\ P(A_j | H_k))}{\sum_{k=1}^{n} P(H_k)\ P(A_j | H_k)}, \quad (2)$$

$$P(H_k | \overline{A}_j) = \frac{P(H_k \overline{A}_j)}{P(\overline{A}_j)} = \frac{P(H_k)\ P(\overline{A}_j | H_k)}{\sum_{k=1}^{n} P(H_k)\ P(\overline{A}_j | H_k)}. \quad (3)$$

В случаях (2) и (3) субъективно выбранными полными группами сравниваемых по степени возможности комбинированных событий являются соответственно множества $\bigcup_{k=1}^{n} H_k A_j$ и $\bigcup_{k=1}^{n} H_k \overline{A}_j$. Это тот случай, когда формула Байеса неприменима, т. к. нарушен принцип сохранения отношений вероятностей — не соблюдается равенство отношений нормированных вероятностей отношениям соответствующих им нормируемых вероятностей:

$$\frac{P(H_k | A_j)}{P(H_k | \overline{A}_j)} = \frac{P(H_k)\ P(A_j | H_k)}{\sum_{k=1}^{n} P(H_k)\ P(A_j | H_k)} \Bigg/ \frac{P(H_k)\ P(\overline{A}_j | H_k)}{\sum_{k=1}^{n} P(H_k)\ P(\overline{A}_j | H_k)} = \frac{P(H_k)\ P(A_j | H_k)}{P(H_k)\ P(\overline{A}_j | H_k)}.$$

Согласно принципу сохранения отношений вероятностей, корректная обработка вероятностей событий осуществима лишь при нормировании с применением одного общего нормирующего делителя, равного сумме всех сравниваемых нормируемых выражений. Поэтому

$$\sum_{k=1}^{n} P(H_k) P(A_j | H_k) + \sum_{k=1}^{n} P(H_k) P(\overline{A}_j | H_k) = \sum_{k=1}^{n} P(H_k)[P(A_j | H_k) + P(H_k)\ P(\overline{A}_j | H_k)] = \sum_{k=1}^{n} P(H_k) = 1.$$

Таким образом, обнаруживается тот факт, что существуют разновидности формулы Байеса, отличающиеся от известных отсутствием нормирующего делителя:

$$P(H_k | A_j) = P(H_k) P(A_j | H_k),\ P(H_k | \overline{A}_j) = P(H_k) P(\overline{A}_j | H_k). \quad (4)$$

При этом соблюдается равенство отношений нормированных вероятностей отношениям соответствующих им нормируемых вероятностей:

$$\frac{P(H_k | A_j)}{P(H_k | \overline{A}_j)} = \frac{P(H_k)\ P(A_j | H_k)}{P(H_k)\ P(\overline{A}_j | H_k)}.$$

На основе субъективного выбора нетрадиционно записываемых полных групп несовместных комбинированных событий можно увеличить количество модификаций формулы Байеса, включающих свидетельства, а также то или иное количество их отрицаний. Например, наиболее полной группе комбинированных событий $\bigcup_{k=1}^{n} \bigcup_{i=1}^{m} H_k E_i \cup \bigcup_{k=1}^{n} \bigcup_{i=1}^{m} H_k \overline{E}_i$ соответствует (с учетом отсутствия нормирующего делителя) модификация формулы Байеса

$$P(H_k | \widetilde{E}_i) = P(H_k) P(\widetilde{E}_i | H_k),$$

где элементарное событие в виде свидетельства $\tilde{E}_i \in \bigcup_{i=1}^{m} E_i \cup \bigcup_{i=1}^{m} \overline{E}_i$ является одним из элементов указанного множества.

При отсутствии отрицаний свидетельств, то есть при $\tilde{E}_i = E_i \in \bigcup_{i=1}^{m} E_i$,

$$P(H_k | E_i) = \frac{P(H_k)\ P(E_i | H_k)}{\sum_{k=1}^{n} P(H_k)\ P(E_i | H_k)}.$$

Таким образом, модификация формулы Байеса, предназначенная для определения сравниваемых по степени возможности условных вероятностей гипотез при одиночных несовместимых свидетельствах выглядит следующим образом. В числителе содержится нормируемая вероятность одного из комбинированных несовместных событий, образующих полную группу, выраженную в виде произведения априорных вероятностей, а в знаменателе — сумма всех

нормируемых вероятностей. При этом соблюдается принцип сохранения отношений вероятностей — и получаемый результат является правильным.

Вероятности гипотез при одиночных совместимых свидетельствах. Формулы Байеса традиционно применяются для определения сравниваемых по степени возможности апостериорных условных вероятностей гипотез H_k ($k = 1,\ldots,n$) при одном из нескольких рассматриваемых совместимых свидетельств E_i ($i = 1,\ldots,m$). В частности (см., например, [5] и [6]), при определении апостериорных условных вероятностей $P(H_1 \mid E_1)$ и $P(H_1 \mid E_2)$ при каждом из двух совместимых свидетельств E_1 и E_2 употребляются формулы вида:

$$P(H_1 \mid E_1) = \frac{P(H_1)\,P(E_1 \mid H_1)}{\sum\limits_{k=1}^{2} P(H_k)\,P(E_1 \mid H_k)} \text{ и } P(H_1 \mid E_2) = \frac{P(H_1)\,P(E_2 \mid H_1)}{\sum\limits_{k=1}^{2} P(H_k)\,P(E_2 \mid H_k)}. \qquad (5)$$

Необходимо учесть, что это еще один случай, когда формула Байеса неприменима. Причем в данном случае должны быть устранены два недостатка:

— проиллюстрированное нормирование вероятностей комбинированных событий некорректно, ввиду принадлежности разным полным группам рассматриваемых событий [10];

— в символических записях комбинированных событий $H_k E_1$ и $H_k E_2$ не находит отражения тот факт, что учитываемые свидетельства E_1 и E_2 являются совместимыми.

Для устранения последнего недостатка может быть использована более развернутая запись комбинированных событий с учетом того, что совместимые свидетельства E_1 и E_2 в одних случаях могут быть несовместными, а в других совместными:

$$H_k E_1 = H_k E_1 E_2 \text{ и } H_k E_2 = H_k \overline{E}_1 E_2 + H_k E_1 E_2,$$

где $\overline{E}_1$ и $\overline{E}_2$ являются свидетельствами, противоположными E_1 и E_2.

Очевидно, что в таких случаях произведение событий $H_k E_1 E_2$ учитывается дважды. Кроме того, оно может быть учтено еще раз отдельно, однако этого не происходит. Дело в том, что в рассматриваемой ситуации на оцениваемую обстановку влияют три вероятных несовместимых комбинированных события: $H_k E_1 \overline{E}_2$, $H_k \overline{E}_1 E_2$ и $H_k E_1 E_2$. При этом для лица, принимающего решение, представляет интерес оценка по степени возможности лишь двух несовместимых комбинированных событий: $H_k E_1 \overline{E}_2$ и $H_k \overline{E}_1 E_2$, что соответствует рассмотрению только одиночных свидетельств.

Таким образом, при построении модификации формулы Байеса для определения апостериорных условных вероятностей гипотез при одиночных совместимых свидетельствах необходимо исходить из следующего. Лицо, принимающее решение, интересует, какое именно элементарное событие, представленное тем или иным свидетельством из числа рассматриваемых, реально произошло в конкретных условиях. Если происходит другое элементарное событие в виде одиночного свидетельства, требуется пересмотр решения, обусловленного результатами сравнительной оценки апостериорных условных вероятностей гипотез с непременным учетом других условий, влияющих на реальную обстановку.

Введем следующее обозначение: $H_k E_i^-$ для одного (и только одного) несовместимого комбинированного события, состоящего в том, что из $m > 1$ рассматриваемых элементарных событий E_i ($i = 1,\ldots,m$) совместно с гипотезой H_k произошло одно элементарное событие E_1 и не произошли другие элементарные события.

В наиболее простом случае рассматриваются два одиночных несовместимых свидетельства. Если подтверждается одно из них, условная вероятность свидетельства в общем виде выражается формулой

$$P(H_k E_i^-) = P(E_i \mid H_k) - P(E_1 E_2 \mid H_k) = P(E_i \mid H_k) - P(E_1 \mid H_k) P(E_2 \mid H_k),\ i = 1,\ldots,2. \qquad (6)$$

В справедливости формулы можно наглядно убедиться (рис. 1).

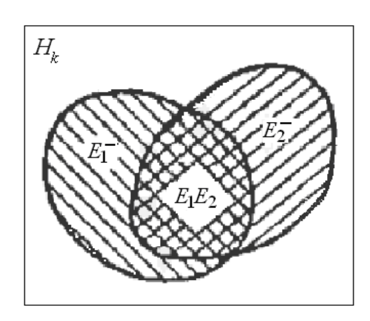


Рис. 1. Геометрическая интерпретация вычисления $P(H_k\, E_i^-)$ при $i = 1,\ldots,2$

При условно независимых свидетельствах

$$P(E_1 E_2 | H_k) = P(E_1 | H_k) P(E_2 | H_k),$$

поэтому с учетом (6)

$$P(H_k\, E_i^-) = P(E_i | H_k) - P(E_1 | H_k) P(E_2 | H_k),\ i = 1,\ldots,2. \qquad (7)$$

Аналогично вероятность $P(H_k E_i^-)$ одного из трех ($i = 1,\ldots,3$) несовместимых событий $H_k E_i^-$ выражается формулой

$$P(H_k\, E_i^-) = P(E_i | H_k) - [\sum_{j=1(j \neq i)}^{3} P(E_i | H_k) P(E_j | H_k)] + P(E_1 E_2 E_3 | H_k).$$

Например, при $i = 1$:

$$P(H_k\, E_1^-) = P(E_1 | H_k) - P(E_1 E_2 | H_k) - P(E_1 E_3 | H_k) + P(E_1 E_2 E_3 | H_k).$$

Справедливость данной формулы наглядно подтверждает геометрическая интерпретация, представленная на рис. 2.

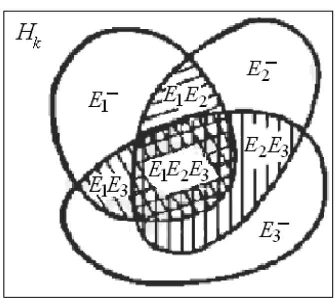


Рис. 2. Геометрическая интерпретация вычисления $P(H_k\, E_i^-)$ при $i = 1,\ldots,3$

Методом математической индукции можно доказать общую формулу для вероятности $P(H_k\, E_i^-)$ при любом количестве свидетельств E_i (i=1,…,m):

$$P(H_k E_i^-) = P(E_i | H_k) - \sum_{j=1(j \neq i)}^{m} P(E_i | H_k)\, P(E_j | H_k) + \sum_{j,l \neq i}^{m} P(E_i | H_k)\, P(E_j | H_k)\, P(E_j | H_k) + \ldots + (-1)^{m-1}. \qquad (8)$$

Используя теорему умножения вероятностей, запишем условную вероятность $P(H_k E_i^-)$ в двух формах:

$$P(H_k\, E_i^-) = P(H_k)\, P(E_i^- | H_k) = P(E_i^-)\, P(H_k | E_i^-),$$

из которых следует, что

$$P(H_k | E_i^-) = \frac{P(H_k E_i^-)}{P(E_i^-)}.$$

С использованием формулы полной вероятности $P(E_i^-) = \sum_{k=1}^{n} P(H_k)\ P(E_i^- | H_k)$ получается, что

$$P(H_k \mid E_i^-) = \frac{P(H_k E_i^-)}{\sum\limits_{k=1}^{n} P(H_k E_i^-)}.$$

Подставив в полученную формулу выражения для $P(H_k E_i^-)$ в виде правой части (8), получим окончательный вид формулы для определения апостериорных условных вероятностей гипотез H_k $(k = 1,\ldots,n)$ при одном из нескольких рассматриваемых несовместимых одиночных свидетельств: $(E_j | H_k)$

$$P(H_k \mid E_i^-) = \frac{P(H_k)[P(E_i|H_k) - \sum\limits_{j=1(j\neq i)}^{m} P(E_i|H_k)\,P(E_j|H_k) + \ldots + (-1)^{m-1} P(\prod\limits_{j=1}^{m} P(E_j|H_k)]}{\sum\limits_{k=1}^{n} P(H_k) \sum\limits_{i=1}^{m} [P(E_i|H_k) - \sum\limits_{j=1(j\neq i)}^{m} P(E_i|H_k)\,P(E_j|H_k) + \ldots + (-1)^{m-1} P(\prod\limits_{j=1}^{m} P(E_j|H_k)]} \quad (9)$$

Сравнительные оценки. Рассматриваются довольно простые, но наглядные примеры, ограничивающиеся анализом вычисляемых апостериорных условных вероятностей одной из двух гипотез при двух одиночных свидетельствах.

1. Вероятности гипотез при несовместимых одиночных свидетельствах. Сравним результаты, получаемые с применением формул Байеса (2) и (3), на примере двух свидетельств $A_j = A$ и $A_j = A$ при исходных данных:

$$P(H_1) = 0{,}7;\ P(H_2) = 0{,}3;\ \ P(A \mid H_1) = 0{,}1;$$

$$P(\overline{A}|H_1) = 0{,}9;\ P(A \mid H_2) = 0{,}6;\ P(\overline{A}|H_2) = 0{,}4.$$

В рассматриваемых примерах с гипотезой H_1 традиционные формулы (2) и (3) приводят к следующим результатам:

$$P(H_1 | A) = \frac{P(H_1)\ P(\mathrm{A}|H_1)}{\sum\limits_{k=1}^{2} P(H_k)\ P(\mathrm{A}|H_k)} = \frac{0.07}{0{,}25} = 0{,}28,$$

$$P(H_1 | \overline{A}) = \frac{P(H_1)\ P(\overline{\mathrm{A}}|H_1)}{\sum\limits_{k=1}^{2} P(H_k)\ P(\overline{\mathrm{A}}|H_k)} = \frac{0.63}{0{,}75} = 0{,}84,$$

а при предлагаемых формулах (4), не имеющих нормирующих делителей:

$$P(H_1 | A) = P(H_1)\ P(A \mid H_1) = 0{,}07;\ \ P(H_1 | \overline{\mathrm{A}}) = P(H_1)\ P(\overline{\mathrm{A}} \mid H_1) = 0{,}63.$$

Таким образом, в случае применения предлагаемых формул отношение нормируемых вероятностей равно отношению нормированных вероятностей:

$$\frac{P(H_1)\ P(\mathrm{A}|H_1)}{P(H_1)\ P(\overline{\mathrm{A}}|H_1)} = \frac{P(\mathrm{A}|H_1)}{P(\overline{\mathrm{A}}|H_1)} = \frac{0{,}07}{0{,}63} = 0{,}11. \quad (10)$$

При использовании известных формул при таком же отношении $\frac{P(H_1)\ P(\mathrm{A}|H_1)}{P(H_1)\ P(\overline{\mathrm{A}}|H_1)} = 0{,}11$ нормируемых вероятностей, указанных в числителях, отношение получаемых нормированных вероятностей:

$$\frac{P(H_1 | A)}{P(H_1 | \overline{A})} = \frac{0{,}28}{0{,}84} = 0{,}33. \quad (11)$$

То есть принцип сохранения отношений вероятностей не соблюдается, и получаются неверные результаты. При этом в случае применения известных формул значение относительного отклонения отношения (11) апостериорных условных вероятностей гипотез от корректных результатов (10) оказывается весьма существенным, так как составляет $\frac{0{,}33 - 0{,}11}{0{,}11} \times 100 = 242\%.$.

2. Вероятности гипотез при совместимых одиночных свидетельствах. Сравним результаты, получаемые с применением формул Байеса (5) и построенной корректной модификации (9), используя следующие исходные данные:

$$P(H_1) = 0{,}7;\ \ P(H_2) = 0{,}3;\ \ P(E_1 \mid H_1) = 0{,}4;$$

$$P(E_2 \mid H_1) = 0{,}8;\ \ P(E_1 \mid H_2) = 0{,}7;\ \ P(E_2 \mid H_2) = 0{,}2.$$

В рассматриваемых примерах с гипотезой H_2 в случае использования традиционных формул (5):

$$P(H_2|E_1) = \frac{P(H_2)\,P(E_1|H_2)}{\sum_{k=1}^{2} P(H_k)\,P(E_1|H_k)} = \frac{0,21}{0,49} = 0,429\ ,$$

$$P(H_2|E_2) = \frac{P(H_2)\,P(E_2|H_2)}{\sum_{k=1}^{2} P(H_k)\,P(E_2|H_k)} = \frac{0,06}{0,62} = 0,097.$$

В случае же применения предлагаемой формулы (9) с учетом (7)

$$P(H_2|E_i^-) = \frac{P(H_2)[P(E_1|H_2) - P(E_1|H_2)\,P(E_2|H_2)]}{\sum_{k=1}^{2} P(H_k)\sum_{i=1}^{2}[P(E_i|H_k) - P(E_1|H_k)\,P(E_2|H_k)]} = \frac{0,168}{0,578} = 0,291,$$

$$P(H_2|E_2^-) = \frac{P(H_2)[P(E_2|H_2) - P(E_1|H_2)\,P(E_2|H_2)]}{\sum_{k=1}^{2} P(H_k)\sum_{i=1}^{2}[P(E_i|H_k) - P(E_1|H_k)\,P(E_2|H_k)]} = \frac{0,018}{0,578} = 0,031.$$

При использовании предлагаемых корректных формул, ввиду одинаковых знаменателей, отношение $\frac{P(H_2)[P(E_1|H_2) - P(E_1|H_2)\,P(E_2|H_2)]}{P(H_2)[P(E_2|H_2) - P(E_1|H_2)\,P(E_2|H_2)]}$ нормируемых вероятностей, указываемых в числителях, равно отношению нормированных вероятностей:

$$\frac{P(\mathrm{H}_2|E_1^-)}{P(\mathrm{H}_2|E_2^-)} = \frac{0,291}{0,031} = 9,387. \qquad (12)$$

То есть принцип сохранения отношений вероятностей соблюдается.

Однако в случае применения известных формул при отношении указанных в числителях нормируемых вероятностей

$$\frac{P(H_2)\,P(E_1|H_2)}{P(H_2)P(E_2|H_2)} = \frac{0,21}{0,06} = 3,5$$

отношение нормированных вероятностей:

$$\frac{P(H_2|E_1)}{P(H_2|E_2)} = \frac{0,429}{0,097} = 4,423. \qquad (13)$$

То есть принцип сохранения отношений вероятностей, как и прежде, не соблюдается. При этом в случае применения известных формул значение относительного отклонения отношения (13) апостериорных условных вероятностей гипотез от корректных результатов (12) также оказывается весьма существенным:

$$\frac{9,387 - 4,423}{9,387} \times 100 = 52,9\%.$$

Заключение. Анализ построения конкретных формульных соотношений, реализующих формулу Байеса и ее модификации, предлагаемые для решения практических задач, позволяют утверждать следующее. Полная группа сравниваемых по степени возможности комбинированных событий может выбираться субъективно лицом, принимающим решение. Данный выбор основывается на учитываемых объективных исходных данных, характерных для типовой обстановки (конкретные виды и количество элементарных событий — оцениваемых гипотез и свидетельств). Представляет практический интерес субъективный выбор других вариантов полной группы сравниваемых по степени возможности комбинированных событий — таким образом обеспечивается существенное разнообразие формульных соотношений при построении нетрадиционных вариантов модификаций формулы Байеса. На этом, в свою очередь, может основываться совершенствование математического обеспечения программной реализации, а также расширение области применения новых формульных соотношений для решения прикладных задач.

Библиографический список

1. Gnedenko, B. V. An elementary introduction to the theory of probability / B. V. Gnedenko, A. Ya. Khinchin. — New York : Dover Publications, 1962. — 144 p.

2. Вентцель, Е. С. Теория вероятностей / Е. С. Вентцель. — 10-е изд., стер. — Москва : Высшая школа, 2006. — 575 с.
3. Андронов. А. М., Теория вероятностей и математическая статистика / А. М. Андронов, Е. А. Копытов, Л. Я. Гринглаз. — Санкт-Петербург : Питер, 2004. — 481 с.
4. Змитрович, А. И. Интеллектуальные информационные системы / А. И. Змитрович. — Минск : ТетраСистемс, 1997. — 496 с.
5. Черноруцкий, И. Г. Методы принятия решений / И. Г. Черноруцкий. — Санкт-Петербург : БХВ-Петербург, 2005. — 416 с.
6. Naylor, C.-M. Build Your Own Expert System / C.-M. Naylor. — Chichester : John Wiley & Sons, 1987. — 289 p.
7. Романов, В. П. Интеллектуальные информационные системы в экономике / В. П. Романов. — 2-е изд., стер. — Москва : Экзамен, 2007. — 496 с.
8. Экономическая эффективность и конкурентоспособность / Д. Ю. Муромцев [и др.]. — Тамбов : Изд-во Тамб. гос. техн. ун-та, 2007.— 96 с.
9. Долгов, А. И. Корректные модификации формулы Байеса для параллельного программирования / А. И. Долгов // Суперкомпьютерные технологии : мат-лы 3-й всерос. науч-техн. конф. — Ростов-на-Дону. — 2014.— Т. 1 — С. 122–126.
10.Долгов, А. И. О корректности модификаций формулы Байеса / А. И. Долгов // Вестник Дон. гос. техн. ун-та. — 2014. — Т. 14, № 3 (78). — С. 13–20.

References

1. Gnedenko, B.V., Khinchin, A.Ya. An elementary introduction to the theory of probability. New York: Dover Publications, 1962, 144 p.
2. Ventsel, E.S. Teoriya veroyatnostey. [Theory of probabilities.] 10th ed., reimpr. Moscow: Vysshaya shkola, 2006, 575 p. (in Russian).
3. Andronov, A.M., Kopytov, E.A., Gringlaz, L.Y. Teoriya veroyatnostey i matematicheskaya statistika. [Theory of probabilities and mathematical statistics.] St.Petersburg: Piter, 2004, 481 p. (in Russian).
4. Zmitrovich, A.I. Intellektual'nye informatsionnye sistemy. [Intelligent information systems.] Minsk: TetraSistems, 1997, 496 p. (in Russian).
5. Chernorutskiy, I.G. Metody prinyatiya resheniy. [Decision-making techniques.] St.Petersburg: BKhV-Peterburg, 2005, 416 p. (in Russian).
6. Naylor, C.-M. Build Your Own Expert System. Chichester: John Wiley & Sons, 1987, 289 p.
7. Romanov, V.P. Intellektual'nye informatsionnye sistemy v ekonomike. [Intelligent information systems in economy.] 2nd ed., reimpr. Moscow: Ekzamen, 2007, 496 p. (in Russian).
8. Muromtsev, D.Y., et al. Ekonomicheskaya effektivnost' i konkurentosposobnost'. [Economic efficiency and competitiveness.] Tambov: Izd-vo Tamb. gos. tekhn. un-ta, 2007, 96 p. (in Russian).
9. Dolgov, A.I. Korrektnye modifikatsii formuly Bayesa dlya parallel'nogo programmirovaniya. [Correct modifications of the Bayesian formula for parallel programming.] Superkomp'yuternye tekhnologii: mat-ly 3-y vseros. nauch-tekhn. konf. [Supercomputer technologies: Proc. III All-Russian Sci.-Tech. Conf.] Rostov-on-Don, 2014, vol. 1, pp. 122–126 (in Russian).
10. Dolgov, A.I. O korrektnosti modifikatsiy formuly Bayesa. [About correctness of Bayes formula modifications.] Vestnik of DSTU, 2014, vol. 14, no. 3 (78), pp. 13–20 (in Russian).

Поступила в редакцию 21.09.2015
Сдана в редакцию 21.09.2015
Запланирована в номер 24.09.2015